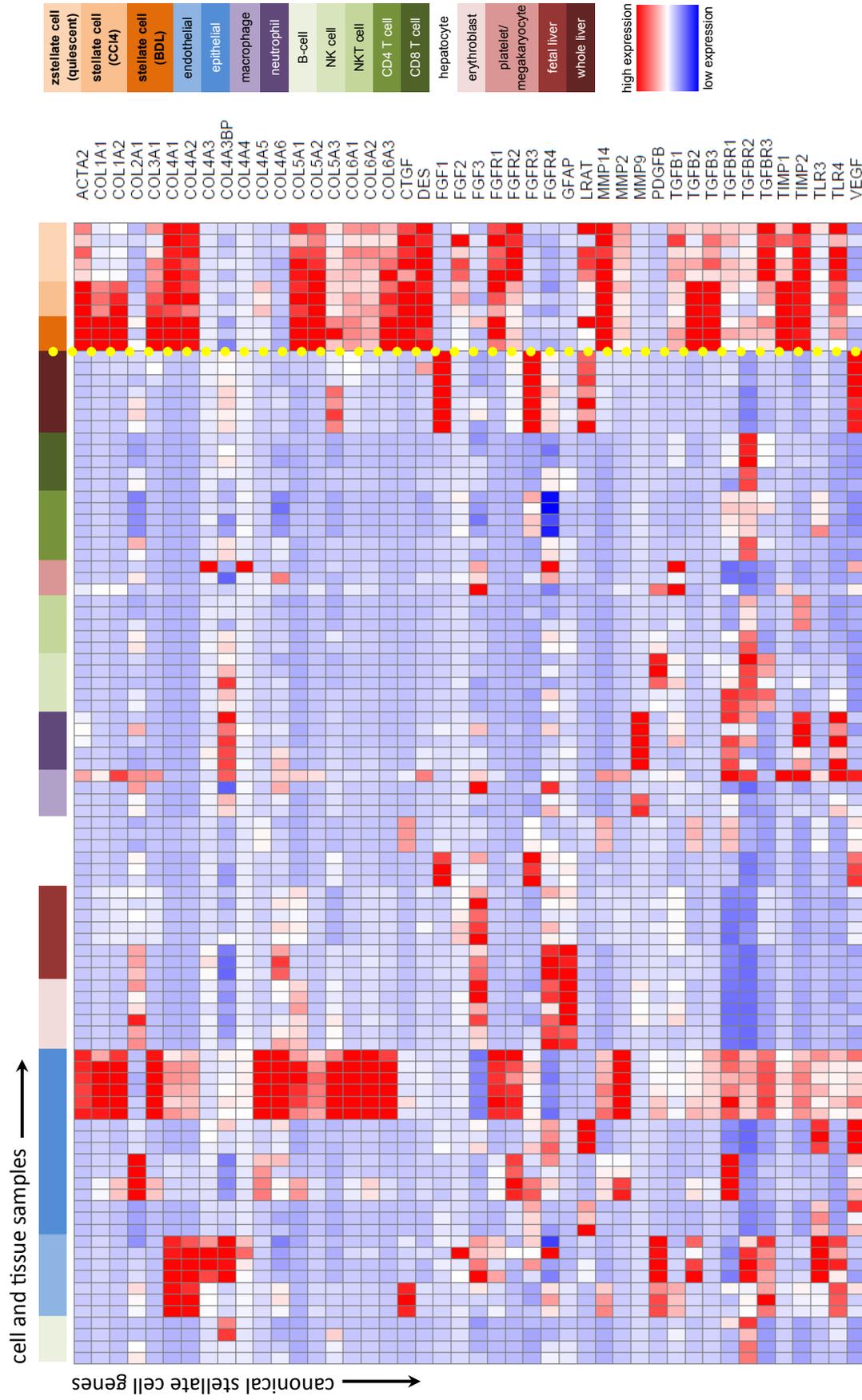


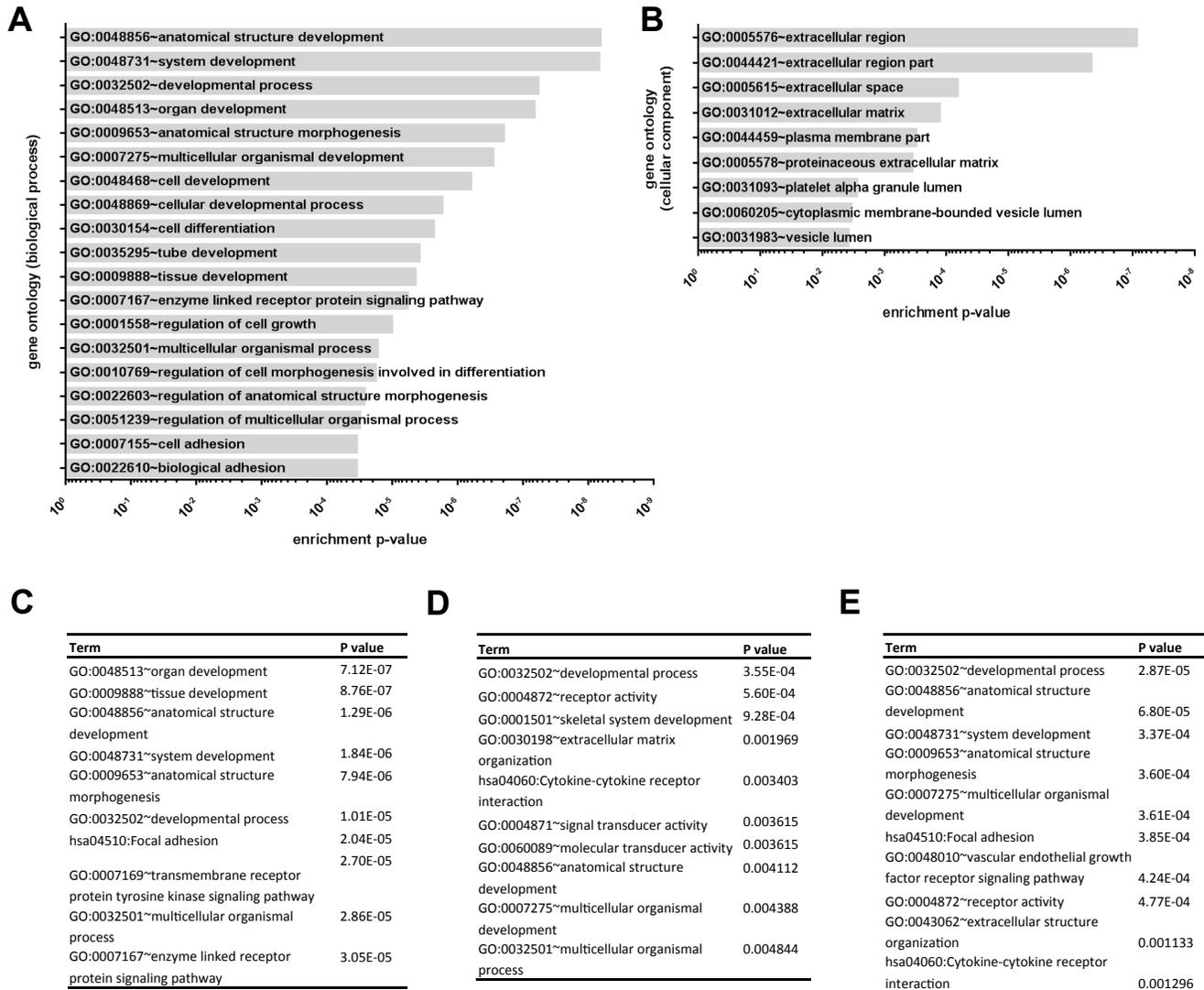


**Supplementary Figure 2: Expression of Canonical Hepatic Stellate Cell Genes across all Tissue and Cell Types**



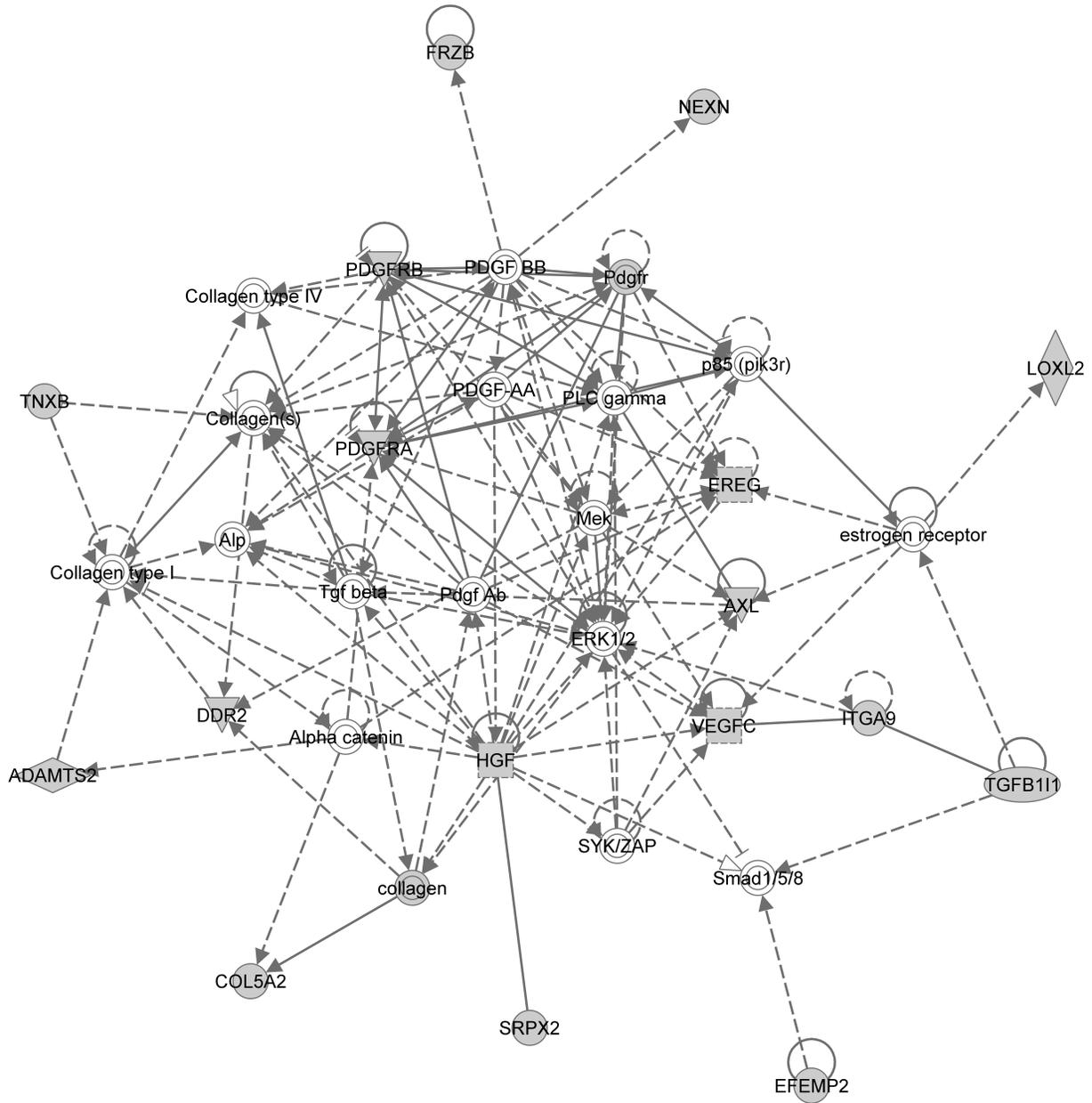
Many canonical hepatic stellate cell factors were enriched in the hepatic stellate cell samples, but were not included in the hepatic stellate cell signature due to concurrent expression in other tissues.

## Supplementary Figure 3: Significantly Enriched Gene Ontologies Within the Hepatic Stellate Cell Signature and Leading Edge Subsets



(A-B) Gene ontologies enriched in the stellate cell signature. (A) shows ontologies describing biological function, where there is strong enrichment of developmentally relevant factors. (B) shows ontologies describing cell location, where there is strong enrichment of extracellular proteins and extracellular matrix components. In (C-E), we characterized the most highly enriched (leading edge) hepatic stellate signature genes from each gene set enrichment analysis in Figure 1. The ten most significantly enriched gene ontologies within each leading edge are listed, comparing normal and cirrhotic patients (C), fibrotic and non-fibrotic NAFLD patients (D), and inflammatory and non-inflammatory NAFLD patients (E).

## Supplementary Figure 4: Interaction Network of Stellate Cell Signature Genes associated with Human Cirrhosis

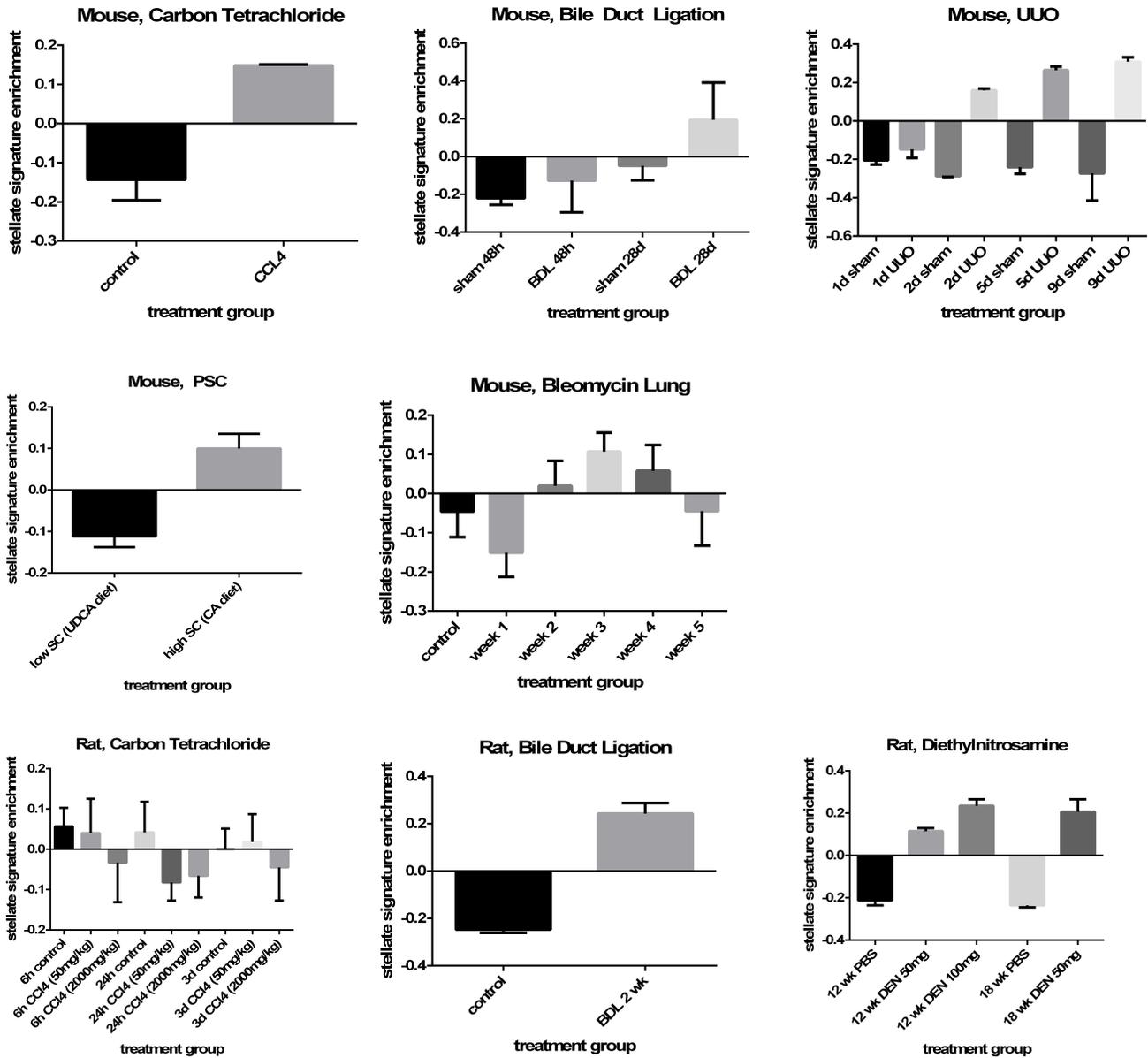


Hepatic stellate cell signature enrichment was assessed using gene set enrichment analysis in a dataset comparing normal and cirrhotic patients. The stellate signature genes most highly correlated with the cirrhotic phenotype were submitted to Ingenuity Pathway Analysis. The highest ranking interaction network from the analysis is shown here.



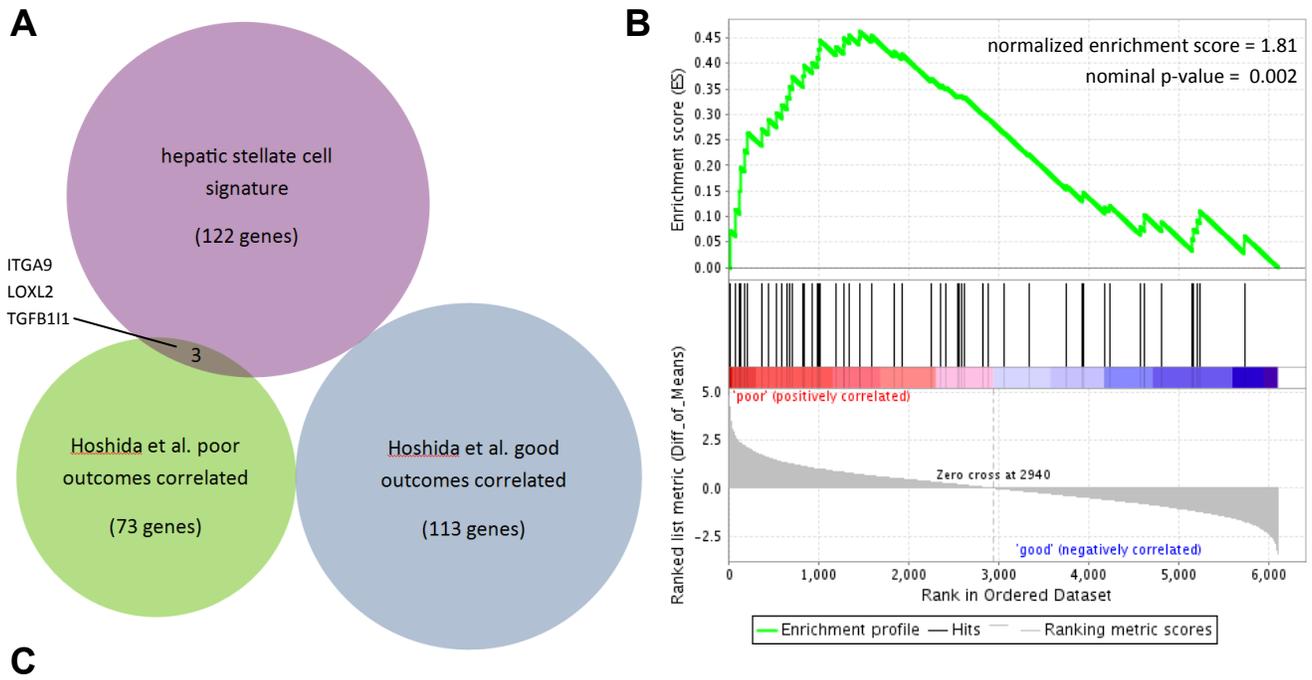


## Supplementary Figure 7: Association of Hepatic Stellate Cell Signature Expression and Rodent Biological Phenotypes



Mean group stellate signature enrichment score is plotted for each condition and dataset indicated. Error bars indicate  $\pm$  SEM. While some interesting trends are present for mouse CCl<sub>4</sub>, mouse BDL, mouse UUO, mouse PSC, rat BDL, and rat DEN, no statistical significance was reached using a two tailed Student's t-test.

## Supplementary Figure 8: Overlap between Hepatic Stellate Cell Signature and Previously Reported Outcomes Signatures (Hoshida et al. 2008 and 2013)



Gene Symbol	Gene Name	Rank in Gene List	Rank Metric Score	Running	Core Enrichment?
ITGA9	ITGA9:integrin, alpha 9	8	3.811203	0.071937	Yes
TGFBI1	TGFBI1:transforming growth factor beta 1 induced transcript 1	70	2.659373	0.112968	Yes
SEMA3A	SEMA3A:sema domain, immunoglobulin domain (ig), short basic domain, secreted, (semaphorin) 3A	118	2.348283	0.150335	Yes
LOXL2	LOXL2:lysyl oxidase-like 2	120	2.341989	0.195187	Yes
SDC3	SDC3:syndecan 3 (N-syndecan)	166	2.182042	0.229689	Yes
TPM1	TPM1:tropomyosin 1 (alpha)	206	2.09128	0.263439	Yes
ADAMTS2	ADAMTS2:ADAM metalloproteinase with thrombospondin type 1 motif, 2	357	1.720657	0.271708	Yes
FRZB	FRZB:frizzled-related protein	429	1.584663	0.290427	Yes
TLN2	TLN2:talín 2	525	1.436981	0.302339	Yes
EDNRA	EDNRA:endothelin receptor type A	584	1.35967	0.318883	Yes
RBMS3	RBMS3:RNA binding motif, single stranded interacting protein	643	1.30307	0.334339	Yes
PDGFRB	PDGFRB:platelet-derived growth factor receptor, beta polypeptide	672	1.267684	0.354076	Yes
DCBLD2	DCBLD2:discoidin, CUB and LCCL domain containing 2	700	1.241085	0.373468	Yes
TNFRSF11B	TNFRSF11B:tumor necrosis factor receptor superfamily, member 11b (osteoprotegerin)	822	1.132176	0.375221	Yes
HGF	HGF:hepatocyte growth factor (hepapoietin A; scatter factor)	833	1.122553	0.395145	Yes
BGN	BGN:biglycan	920	1.045854	0.401026	Yes
COL5A2	COL5A2:collagen, type V, alpha 2	978	1.006057	0.410939	Yes
BDNF	BDNF:brain-derived neurotrophic factor	993	0.999173	0.42783	Yes
EDNRB	EDNRB:endothelin receptor type B	1007	0.994661	0.4448	Yes
PDGFRA	PDGFRA:platelet-derived growth factor receptor, alpha polypeptide	1186	0.868962	0.432067	Yes
TAGLN	TAGLN:transgelin	1269	0.818863	0.434247	Yes
TNXB	TNXB:tenascin XB	1275	0.815062	0.449087	Yes
ANGPTL2	ANGPTL2:angiopoietin-like 2	1332	0.773426	0.454693	Yes
TGFB3	TGFB3:transforming growth factor, beta 3	1444	0.710505	0.449995	Yes
RGS7	RGS7:regulator of G-protein signalling 7	1455	0.705168	0.461896	Yes

(A) shows the gene overlap between the hepatic stellate cell and previously reported prognostic signature. Only 3 genes are the same, indicating that the stellate cell and prognostic signatures are substantially different. (B) shows the position of stellate signature genes on the resected HCC expression dataset, ranked by Cox regression score for survival. The stellate signature genes with the highest correlation to poor outcome are shown in (C).

# Supplementary Table 1: Datasets Used to Derive the Hepatic Stellate Cell Signature

Sample accession	Description	Category	Sample accession	Description	Category
GSM686644	B220(+) B lymphocytes	B cell	GSM1301661	bone marrow neutrophils (Gr-1 high CD48 neg)	neutrophil
GSM686645	B220(+) B lymphocytes	B cell	GSM1301662	bone marrow neutrophils (Gr-1 high CD48 neg)	neutrophil
GSM1061907	B6 T1 B cells	B cell	GSM344315	Normal BM neutrophils, c-kit- Gr1+ Mac1+	neutrophil
GSM1061908	B6 T1 B cells	B cell	GSM344316	Normal BM neutrophils, c-kit- Gr1+ Mac1+	neutrophil
GSM690763	Liver sinusoidal endothelial cells isolated from MxCre- Fk1 L/L Fk2 L/L Afx L/L mouse	endothelial	GSM344318	Normal BM neutrophils, c-kit- Gr1+ Mac1+	neutrophil
GSM690764	Liver sinusoidal endothelial cells isolated from MxCre- Fk1 L/L Fk2 L/L Afx L/L mouse	endothelial	GSM686650	NK1.1+ NK cells	NK
GSM690765	Liver sinusoidal endothelial cells isolated from MxCre- Fk1 L/L Fk2 L/L Afx L/L mouse	endothelial	GSM686651	NK1.1+ NK cells	NK
GSM547760	Mouse, colon, normal, blood endothelial cells	endothelial	GSM1214478	NK cells were enriched by negative selection using NK cell isolation kit (R&D) and then FACS sorted to obtain DX5+ cells	NK
GSM547762	Mouse, colon, normal, blood endothelial cells	endothelial	GSM1214479	NK cells were enriched by negative selection using NK cell isolation kit (R&D) and then FACS sorted to obtain DX5+ cells	NK
GSM547764	Mouse, colon, normal, blood endothelial cells	endothelial	GSM1214480	NK cells were enriched by negative selection using NK cell isolation kit (R&D) and then FACS sorted to obtain DX5+ cells	NK
GSM547766	Mouse, colon, normal, blood endothelial cells	endothelial	GSM686652	CD3e(+)NK1.1(+) NKT cells	NKT
GSM765922	c57 bl/6 intestinal epithelial	epithelial	GSM686653	CD3e(+)NK1.1(+) NKT cells	NKT
GSM765923	c57 bl/6 intestinal epithelial	epithelial	GSM1232678	stage 2 iNKT	NKT
GSM765924	c57 bl/6 intestinal epithelial	epithelial	GSM1232679	stage 2 iNKT	NKT
GSM598999	microdissected uterine bud stalks from E12.5 kidneys	epithelial	GSM1232680	stage 2 iNKT	NKT
GSM599000	microdissected uterine bud stalks from E12.5 kidneys	epithelial	GSM389821	megakaryocytes from wild type fetal livers E14.5 (CD41+)	platelet/megakaryocyte
GSM599001	microdissected uterine bud tips from E12.5 kidneys	epithelial	GSM1071644	washed platelets from Smad4 f/f mice	platelet/megakaryocyte
GSM190795	intestinal epithelial cells	epithelial	GSM361592	megakaryocytes SCL fl/fl	platelet/megakaryocyte
GSM190796	intestinal epithelial cells	epithelial	GSM852341	Isolated primary hepatic stellate cells, macrophage depleted	stellate cell, BDL
GSM190797	intestinal epithelial cells	epithelial	GSM852342	Isolated primary hepatic stellate cells, macrophage depleted	stellate cell, BDL
GSM658893	FVB/N MMTV-rtTA mammary epithelial cells, Doxycycline treated 7 days	epithelial	GSM852343	Isolated primary hepatic stellate cells, macrophage depleted	stellate cell, BDL
GSM658894	FVB/N MMTV-rtTA mammary epithelial cells, Doxycycline treated 7 days	epithelial	GSM852344	Isolated primary hepatic stellate cells, macrophage depleted	stellate cell, CCl <sub>4</sub>
GSM658895	FVB/N MMTV-rtTA mammary epithelial cells, Doxycycline treated 7 days	epithelial	GSM852345	Isolated primary hepatic stellate cells, macrophage depleted	stellate cell, CCl <sub>4</sub>
GSM658896	FVB/N MMTV-rtTA mammary epithelial cells, Doxycycline treated 7 days	epithelial	GSM852346	Isolated primary hepatic stellate cells, macrophage depleted	stellate cell, CCl <sub>4</sub>
GSM658897	FVB/N MMTV-rtTA mammary epithelial cells, Doxycycline treated 7 days	epithelial	GSM852330	Isolated primary hepatic stellate cells, macrophage depleted	stellate cell, quiescent
GSM658898	FVB/N MMTV-rtTA mammary epithelial cells, Doxycycline treated 7 days	epithelial	GSM852331	Isolated primary hepatic stellate cells, macrophage depleted	stellate cell, quiescent
GSM686654	Ter119(+) erythroblasts	erythroblast	GSM852332	Isolated primary hepatic stellate cells, macrophage depleted	stellate cell, quiescent
GSM686655	Ter119(+) erythroblasts	erythroblast	GSM852333	Isolated primary hepatic stellate cells, macrophage depleted	stellate cell, quiescent
GSM216494	Sorted erythroblasts, Sod2+/-	erythroblast	GSM852334	Isolated primary hepatic stellate cells, macrophage depleted	stellate cell, quiescent
GSM216495	Sorted erythroblasts, Sod2+/-	erythroblast	GSM686646	CD4(+) T lymphocytes	CD4 T cell
GSM216496	Sorted erythroblasts, Sod2+/-	erythroblast	GSM686647	CD4(+) T lymphocytes	CD4 T cell
GSM216497	Sorted erythroblasts, Sod2+/-	erythroblast	GSM555381	CD4+ T cells from B6 mice	CD4 T cell
GSM298115	E14.5 fetal liver Ter119 + cells	fetal liver	GSM555382	CD4+ T cells from B6 mice	CD4 T cell
GSM298116	E14.5 fetal liver Ter119 + cells	fetal liver	GSM555383	CD4+ T cells from B6 mice	CD4 T cell
GSM298117	E14.5 fetal liver Ter119 + cells	fetal liver	GSM555384	CD4+ T cells from B6 mice	CD4 T cell
GSM378250	Embryonic day 14.5 wild type (WT) mouse fetal liver from B6/129 mixed strain	fetal liver	GSM686648	CD8(+) T lymphocytes	CD8 T cell
GSM378251	Embryonic day 14.5 wild type (WT) mouse fetal liver from B6/129 mixed strain	fetal liver	GSM686649	CD8(+) T lymphocytes	CD8 T cell
GSM378252	Embryonic day 14.5 wild type (WT) mouse fetal liver from B6/129 mixed strain	fetal liver	GSM1081398	Quiescent mouse CD8 T cells purified from WT control mouse spleen	CD8 T cell
GSM378253	Embryonic day 14.5 wild type (WT) mouse fetal liver from B6/129 mixed strain	fetal liver	GSM1081399	Quiescent mouse CD8 T cells purified from WT control mouse spleen	CD8 T cell
GSM378254	Embryonic day 14.5 wild type (WT) mouse fetal liver from B6/129 mixed strain	fetal liver	GSM1081400	Quiescent mouse CD8 T cells purified from WT control mouse spleen	CD8 T cell
GSM1281320	freshly isolated primary mouse hepatocytes	hepatocyte	GSM591473	normal liver from C57BL/6 mice fed a low fat diet	whole liver
GSM1281321	freshly isolated primary mouse hepatocytes	hepatocyte	GSM591475	normal liver from C57BL/6 mice fed a low fat diet	whole liver
GSM1281322	freshly isolated primary mouse hepatocytes	hepatocyte	GSM591477	normal liver from C57BL/6 mice fed a low fat diet	whole liver
GSM525384	primary mouse hepatocytes	hepatocyte	GSM591480	normal liver from C57BL/6 mice fed a low fat diet	whole liver
GSM525383	primary mouse hepatocytes	hepatocyte	GSM602665	Six week old male C57BL/6 mouse liver	whole liver
GSM525382	primary mouse hepatocytes	hepatocyte	GSM602666	Six week old male C57BL/6 mouse liver	whole liver
GSM686658	Mac-1(+) monocytes/macrophages	macrophage	GSM602667	Six week old male C57BL/6 mouse liver	whole liver
GSM686659	Mac-1(+) monocytes/macrophages	macrophage			
GSM571897	fetal liver macrophage Mac-1+F4/80+	macrophage			
GSM1129665	cultured peritoneal macrophages	macrophage			

GSM: NCBI Gene Expression Omnibus sample accession number (<http://www.ncbi.nlm.nih.gov/geo/>)

## Supplementary Table 2: Gene Ontology Terms Enriched in the Hepatic Stellate Cell Signature

GO Term (Molecular Function)	Count	% of term	Genes	P-Value	Bonferroni	FDR
GO:0048856--anatomical structure development	44	3.93	BMP10, TNC, TBX20, PTH1R, TGFB3, CDH2, TPM1, EDNRA, EDNRB, TNFRSF11B, BDNF, HAND2, LHX2, CCBE1, SEMA3D, AHNAK2, ROBO2, HHIP, SEMA3A, CRYAB, AXL, TEAD1, HGF, FRZB, COL5A2, NTN1, TNS3, VEGFC, BGN, SEMA6D, EREG, TAGLN, PRICKLE2, PDGFRA, PDGFRB, GDF10, ZFPM2, RELN, ANTXR1, TGFB11, ADAMTS2, BMP5, NTM, NGF	6.43E-09	7.61E-06	1.04E-05
GO:0048731--system development	42	3.75	BMP10, TNC, TBX20, PTH1R, TGFB3, CDH2, TPM1, EDNRA, EDNRB, TNFRSF11B, BDNF, HAND2, LHX2, CCBE1, SEMA3D, AHNAK2, ROBO2, SEMA3A, HHIP, CRYAB, AXL, TEAD1, HGF, FRZB, COL5A2, NTN1, TNS3, VEGFC, BGN, SEMA6D, EREG, TAGLN, PDGFRA, PDGFRB, GDF10, ZFPM2, RELN, TGFB11, ADAMTS2, BMP5, NTM, NGF	6.74E-09	7.99E-06	1.09E-05
GO:0032502--developmental process	48	4.29	BMP10, TNC, TBX20, PTH1R, TGFB3, CDH2, TPM1, FOXS1, EDNRA, EDNRB, BDNF, TNFRSF11B, HAND2, LHX2, CCBE1, SEMA3D, AHNAK2, ROBO2, HHIP, SEMA3A, LOXL2, ANGPTL2, CRYAB, AXL, TEAD1, PDE3A, HGF, FRZB, COL5A2, NTN1, VEGFC, TNS3, BGN, SEMA6D, EREG, TAGLN, PRICKLE2, PDGFRA, PDGFRB, GDF10, ZFPM2, RELN, ANTXR1, TGFB11, ADAMTS2, BMP5, NTM, NGF	5.82E-08	6.90E-05	9.39E-05
GO:0048513--organ development	34	3.04	BMP10, TNC, TBX20, PTH1R, TGFB3, CDH2, TPM1, EDNRA, EDNRB, BDNF, HAND2, LHX2, CCBE1, AHNAK2, ROBO2, HHIP, CRYAB, AXL, TEAD1, HGF, NTN1, COL5A2, TNS3, VEGFC, BGN, EREG, TAGLN, PDGFRA, PDGFRB, ZFPM2, RELN, TGFB11, ADAMTS2, BMP5	6.66E-08	7.89E-05	1.07E-04
GO:0009653--anatomical structure morphogenesis	27	2.41	BMP10, TBX20, TGFB3, CDH2, TPM1, EDNRA, BDNF, HAND2, LHX2, CCBE1, ROBO2, SEMA3A, HHIP, AXL, HGF, NTN1, COL5A2, VEGFC, BGN, EREG, PRICKLE2, PDGFRA, PDGFRB, ZFPM2, RELN, ANTXR1, TGFB11	1.97E-07	2.33E-04	3.18E-04
GO:0007275--multicellular organismal development	44	3.93	BMP10, TNC, TBX20, PTH1R, TGFB3, CDH2, TPM1, FOXS1, EDNRA, EDNRB, TNFRSF11B, BDNF, HAND2, LHX2, CCBE1, SEMA3D, AHNAK2, ROBO2, HHIP, SEMA3A, ANGPTL2, CRYAB, AXL, TEAD1, HGF, FRZB, COL5A2, NTN1, TNS3, VEGFC, BGN, SEMA6D, EREG, TAGLN, PDGFRA, PDGFRB, GDF10, ZFPM2, RELN, TGFB11, ADAMTS2, BMP5, NTM, NGF	2.83E-07	3.36E-04	4.57E-04
GO:0048468--cell development	19	1.70	BMP10, TNC, PTH1R, PDE3A, HGF, NTN1, TPM1, EDNRA, EDNRB, BDNF, EREG, HAND2, LHX2, ROBO2, RELN, ANTXR1, SEMA3A, NTM, NGF	6.21E-07	7.35E-04	1.00E-03
GO:0048869--cellular developmental process	31	2.77	BMP10, TNC, PTH1R, CDH2, TPM1, EDNRA, EDNRB, BDNF, HAND2, LHX2, SEMA3D, AHNAK2, ROBO2, SEMA3A, HHIP, PDE3A, HGF, FRZB, NTN1, VEGFC, SEMA6D, EREG, PRICKLE2, PDGFRA, ZFPM2, RELN, ANTXR1, TGFB11, NTM, BMP5, NGF	1.71E-06	2.02E-03	2.76E-03
GO:0030154--cell differentiation	30	2.68	BMP10, TNC, PTH1R, CDH2, TPM1, EDNRA, EDNRB, BDNF, HAND2, LHX2, SEMA3D, AHNAK2, ROBO2, SEMA3A, HHIP, PDE3A, HGF, FRZB, NTN1, VEGFC, EREG, SEMA6D, PDGFRA, PDGFRB, ZFPM2, RELN, ANTXR1, TGFB11, NTM, BMP5, NGF	2.34E-06	2.76E-03	3.77E-03
GO:0035295--tube development	11	0.98	EDNRA, TNS3, BDNF, HAND2, TBX20, PDGFRA, TGFB3, ZFPM2, ROBO2, HHIP, ADAMTS2	3.84E-06	4.55E-03	6.21E-03
GO:0009888--tissue development	18	1.61	BMP10, TNC, PTH1R, HGF, TPM1, COL5A2, EDNRA, EDNRB, VEGFC, EREG, HAND2, LHX2, PDGFRB, AHNAK2, ZFPM2, TGFB11, ADAMTS2, BMP5	4.44E-06	5.25E-03	7.17E-03
GO:0007167--enzyme linked receptor protein signaling pathway	13	1.16	BMP10, GDF2, AXL, TGFB3, HGF, DDR2, VEGFC, EREG, PDGFRA, GDF10, PDGFRB, TGFB11, NGF	5.81E-06	6.86E-03	9.38E-03
GO:0001558--regulation of cell growth	10	0.89	DCBLD2, RERG, VEGFB, BMP10, GDF2, CRYAB, SEMA3A, NTN1, DDR2, NGF	1.02E-05	1.20E-02	1.65E-02
GO:0032501--multicellular organismal process	52	4.65	BMP10, TBX20, TGFB3, FOXS1, EDNRA, EDNRB, TNFRSF11B, BDNF, CCBE1, SEMA3D, GUCY1A3, ROBO2, SEMA3A, HHIP, KHDRBS3, GPR176, CRYAB, MMP10, VEGFC, TNS3, BGN, TAGLN, EREG, PDGFRA, PDGFRB, ZFPM2, RELN, TGFB11, ADAMTS2, NGF, TNC, PTH1R, CDH2, TPM1, HAND2, KIRREL, LHX2, AHNAK2, ANGPTL2, TNXB, EFEMP2, AXL, TEAD1, PDE3A, HGF, FRZB, COL5A2, NTN1, SEMA6D, GDF10, BMP5, NTM	1.67E-05	1.95E-02	2.69E-02
GO:0010769--regulation of cell morphogenesis involved in differentiation	7	0.63	TGFB3, ROBO2, TGFB11, SEMA3A, CDH2, NTN1, NGF	1.82E-05	2.13E-02	2.93E-02
GO:0022603--regulation of anatomical structure morphogenesis	10	0.89	BDNF, TNFRSF11B, TGFB3, ROBO2, TGFB11, HGF, SEMA3A, CDH2, NTN1, NGF	2.68E-05	3.13E-02	4.33E-02
GO:0051239--regulation of multicellular organismal process	20	1.79	BMP10, TGFB3, AFAP1L2, CDH2, HGF, NTN1, TPM1, FOXS1, EDNRB, VEGFC, TNFRSF11B, BDNF, EREG, HSPB7, GDF10, GUCY1A3, ROBO2, SEMA3A, TGFB11, NGF	3.12E-05	3.63E-02	5.04E-02
GO:0007155--cell adhesion	17	1.52	DCBLD2, SVEP1, TNXB, TLN2, TNC, CDH2, PCDH7, DDR2, SDC3, SCARF2, ITGA9, RELN, ROBO2, ANTXR1, TGFB11, LOXL2, NTM	3.49E-05	4.05E-02	5.63E-02
GO:0022610--biological adhesion	17	1.52	DCBLD2, SVEP1, TNXB, TLN2, TNC, CDH2, PCDH7, DDR2, SDC3, SCARF2, ITGA9, RELN, ROBO2, ANTXR1, TGFB11, LOXL2, NTM	3.55E-05	4.12E-02	5.73E-02
GO Term (Cellular Component)	Count	%	Genes	P-Value	Bonferroni	FDR
GO:0005576--extracellular region	36	3.217	BMP10, GDF2, MASP1, TNC, PAMR1, TGFB3, TNFRSF11B, BDNF, RSPO3, C1QTNF2, CCBE1, SEMA3D, SEMA3A, HHIP, LOXL2, ANGPTL2, P15, TNXB, SVEP1, EFEMP2, HGF, FRZB, COL5A2, NTN1, VEGFB, MMP10, TSLP, VEGFC, BGN, EREG, SRPX2, GDF10, RELN, ADAMTS2, BMP5, NGF	8.62E-08	1.37E-05	1.04E-04
GO:0044421--extracellular region part	23	2.055	BMP10, GDF2, TNXB, MASP1, EFEMP2, TNC, TGFB3, NTN1, COL5A2, MMP10, TSLP, VEGFC, TNFRSF11B, BGN, EREG, C1QTNF2, GDF10, RELN, LOXL2, ANGPTL2, ADAMTS2, BMP5, NGF	4.69E-07	7.45E-05	5.65E-04
GO:0005615--extracellular space	16	1.430	BMP10, GDF2, TNXB, MASP1, TGFB3, MMP10, VEGFC, TSLP, EREG, C1QTNF2, GDF10, RELN, LOXL2, ANGPTL2, BMP5, NGF	6.57E-05	1.04E-02	7.92E-02
GO:0031012--extracellular matrix	11	0.983	MMP10, TNFRSF11B, BGN, TNXB, TNC, EFEMP2, TGFB3, RELN, COL5A2, NTN1, ADAMTS2	1.29E-04	2.02E-02	1.55E-01
GO:0044459--plasma membrane part	30	2.681	DCBLD2, IL1R1, TLN2, GNAI1, PTH1R, TGFB3, SYT9, CDH2, DDR2, AMOTL2, TPM1, EDNRA, EDNRB, ROBO2, HHIP, GPR176, AXL, PCDH7, NEXN, TNS3, ITGA9, ABCC9, EREG, PRICKLE2, PDGFRA, RGS7, PDGFRB, ANTXR1, TGFB11, STEAP1	3.10E-04	4.81E-02	3.74E-01
GO:0005578--proteinaceous extracellular matrix	10	0.894	MMP10, TNFRSF11B, BGN, TNXB, TNC, EFEMP2, RELN, COL5A2, NTN1, ADAMTS2	3.53E-04	5.45E-02	4.25E-01
GO:0031093--platelet alpha granule lumen	4	0.357	VEGFB, VEGFC, TGFB3, HGF	2.76E-03	3.56E-01	3.28E+00
GO:0060205--cytoplasmic membrane-bounded vesicle lumen	4	0.357	VEGFB, VEGFC, TGFB3, HGF	3.38E-03	4.16E-01	4.00E+00
GO:0031983--vesicle lumen	4	0.357	VEGFB, VEGFC, TGFB3, HGF	3.83E-03	4.57E-01	4.53E+00
GO Term (Molecular Function)	Count	%	Genes	P-Value	Bonferroni	FDR
GO:0008083--growth factor activity	11	0.98	VEGFB, VEGFC, BMP10, GDF2, BDNF, EREG, TGFB3, GDF10, HGF, BMP5, NGF	2.86E-07	7.47E-05	3.74E-04
GO:0005102--receptor binding	23	2.06	BMP10, IL1R1, GDF2, TNXB, TNC, TGFB3, HGF, RERG, VEGFB, TSLP, EDNRB, VEGFC, TNFRSF11B, BDNF, EREG, C1QTNF2, PDGFRA, GDF10, PDGFRB, TGFB11, ANGPTL2, BMP5, NGF	3.83E-07	9.99E-05	5.01E-04
GO:0005515--protein binding	80	7.15	BMP10, GDF2, MASP1, GPR124, TLN2, TBX20, TGFB3, SYT9, ANKRD1, NAP1L5, DDR2, AMOTL2, SDC3, EDNRA, EDNRB, BDNF, TNFRSF11B, BOK, PLOD2, GUCY1A3, ROBO2, LMOD1, HHIP, PPP1R14A, CDK14, RPP25, KHDRBS3, CRYAB, LRRRC27, PCDH7, NEXN, MSC, VEGFB, TNS3, VEGFC, BGN, TAGLN, EREG, HSPB7, PDGFRA, PDGFRB, ZFPM2, RELN, TGFB11, FKBP10, IRAK1BP1, NGF, DCBLD2, IL1R1, GNAI1, TNC, AFAP1L2, CDH2, TPM1, SCARF2, HAND2, PDE1A, C1QTNF2, AHNAK2, ANGPTL2, SVEP1, TNXB, LRRN4, EFEMP2, LMCD1, AXL, TEAD1, HGF, FRZB, NTN1, COL5A2, RERG, TSLP, ITGA9, ABCC9, RGS7, GDF10, ANTXR1, NTM, BMP5	1.35E-04	3.45E-02	1.76E-01
Term (KEGG Pathways)	Count	%	Genes	P-Value	Bonferroni	FDR
hsa04510:Focal adhesion	11	0.98	VEGFB, ITGA9, VEGFC, TNXB, TLN2, TNC, PDGFRA, PDGFRB, RELN, HGF, COL5A2	5.70E-06	3.08E-04	5.58E-03
hsa04512:ECM-receptor interaction	6	0.54	ITGA9, TNXB, TNC, RELN, COL5A2, SDC3	7.14E-04	3.79E-02	6.98E-01
hsa04060:Cytokine-cytokine receptor interaction	9	0.80	VEGFB, VEGFC, TSLP, IL1R1, TNFRSF11B, PDGFRA, TGFB3, PDGFRB, HGF	1.57E-03	8.12E-02	1.52E+00

### Supplementary Table 3: Transcriptome Datasets Used for Hepatic Stellate Cell Signature Validation

Organism	Disease / Model	GEO accession
Human	Non-alcoholic fatty liver disease	GSE48452
Mouse	Cirrhosis	GSE6764
	Carbon tetrachloride	GSE27640
	Bile duct ligation	GSE40041
	Primary sclerosing cholangitis	GSE11507
Rat	Unilateral urethral obstruction	GSE36496
	Lung bleomycin	GSE37635
	Carbon tetrachloride	GSE32891
	Bile duct ligation	GSE13747
	Diethylnitrosamine	GSE19057

GSE: NCBI Gene Expression Omnibus sample accession number  
(<http://www.ncbi.nlm.nih.gov/geo/>)

## Supplementary Table 4: List of Gene Ontology and LOCATE Categories Used for Cell Surface Marker Discovery

Gene Ontology Cellular Component	LOCATE
ADHERENS_JUNCTION	GPI anchored membrane protein
ANCHORED_TO_PLASMA_MEMBRANE	Multi-span membrane protein
APICAL_PLASMA_MEMBRANE	Type I membrane protein
APICOLATERAL_PLASMA_MEMBRANE	Type II membrane protein
BASAL_LAMINA	<a href="http://locate.imb.uq.edu.au/">locate.imb.uq.edu.au/</a>
BASEMENT_MEMBRANE	
BASOLATERAL_PLASMA_MEMBRANE	
BRUSH_BORDER	
CELL_JUNCTION	
CELL_MATRIX_JUNCTION	
CELL_PROJECTION	
CELL_PROJECTION_PART	
CELL_SUBSTRATE_ADHERENS_JUNCTION	
CELL_SURFACE	
DYSTROPHIN_ASSOCIATED_GLYCOPROTEIN_COMPLEX	
EXTERNAL_SIDE_OF_PLASMA_MEMBRANE	
EXTRACELLULAR_MATRIX	
EXTRINSIC_TO_PLASMA_MEMBRANE	
FOCAL_ADHESION	
INTEGRAL_TO_PLASMA_MEMBRANE	
INTEGRIN_COMPLEX	
INTRINSIC_TO_PLASMA_MEMBRANE	
LIPID_RAFT	
PLASMA_MEMBRANE	
PLASMA_MEMBRANE_PART	
PROTEINACEOUS_EXTRACELLULAR_MATRIX	
RECEPTOR_COMPLEX	
SYNAPSE	
SYNAPSE_PART	
TIGHT_JUNCTION	

[geneontology.org/page/cellular-component-ontology-guidelines](http://geneontology.org/page/cellular-component-ontology-guidelines)

**Supplementary Table 5: Prognostic Association of Clinical Variables, Clinical Indices, and Hepatic Stellate Cell Signature Expression with Survival (Prognostic Index Derivation Set, HCV-related Cirrhosis)**

	Univariable Regression	Multivariable Regression
n	216	216
Bilirubin		
(>1.0 mg/dL)		
hazard ratio	3.1	3.8
HR, 95% CI	1.8-5.4	1.5-9.6
p-value	<0.001	0.004
Platelet count		
(<100k / mm <sup>3</sup> )		
hazard ratio	3.1	2.4
HR, 95% CI	1.8-5.5	1.2-4.8
p-value	<0.001	0.015
Hepatic stellate cell signature expression		
(>75th percentile)		
hazard ratio	2.5	3.0
HR, 95% CI	1.5-4.1	1.5-5.7
p-value	<0.001	0.001
FIB4 > 3.25		
hazard ratio	2.0	-
95%CI	0.76-5.1	-
p-value	0.16	-
Albu < 4.1 (median)		
hazard ratio	1.9	1.9
95%CI	1.1-3.1	0.97-3.8
p-value	0.014	0.06
MELD >6.4 (median)		
hazard ratio	2.3	1.1
95%CI	1.2-4.3	0.52-2.3
p-value	0.009	0.83

## Supplementary Table 6: Comparison of Prognostic Indices with and without Hepatic Stellate Cell Signature Expression

Cohort	Outcome	c-statistic for clinical model (including bilirubin and platelets)	c-statistic for stellate signature model (including bilirubin, platelets and stellate signature)	p-value for model comparison (likelihood ratio test)
Prognostic index training set (HCV cohort)	Death	0.66 (0.59-0.74)	0.70 (0.62-0.78)	0.0015
	Child progression	0.68 (0.61-0.75)	0.70 (0.63-0.78)	0.0043
	Decompensation	0.61 (0.54-0.68)	0.62 (0.55-0.69)	0.075
	Development of HCC	0.63 (0.55-0.70)	0.63 (0.56-0.71)	0.75
Prognostic index validation set (HCC cohort)	Death	0.58 (0.48-0.68)	0.62 (0.51-0.72)	0.026
	HCC recurrence	0.51 (0.44-0.59)	0.54 (0.46-0.62)	0.11

c-statistic was calculated using multivariable Cox regression model adjusted for bilirubin  $\geq 1$  mg/dl and platelet count  $<100,000/\text{mm}^3$ .

Discrimination of the models was assessed by the c-statistic and improvement of overall model fit between models was assessed by the likelihood ratio test.

## Supplementary Table 7: Correlation of Hepatic Stellate Cell Signature Subsets with Cirrhosis, Fibrosis, and Inflammation

Signature Subset	Correlation with Cirrhosis		Correlation with Fibrosis in NASH		Correlation with Inflammation in NASH	
	NES	p-value	NES	p-value	NES	p-value
Whole hepatic stellate cell signature	-2.08	< 0.001	-1.97	< 0.001	-1.49	0.009
Genes significantly higher in quiescent stellate cells*	-1.92	< 0.001	-1.93	< 0.001	-1.36	0.051
Genes significantly higher in activated stellate cells*	-2.19	< 0.001	-1.93	< 0.001	-1.47	0.028
Genes with no substantial change in expression between activated and quiescent HSC	-1.94	< 0.001	-1.88	0.003	-1.32	0.1

NES: normalized enrichment score in gene set enrichment analysis

\*significantly higher as defined by a p-value < 0.05 in two tailed Student's t-test between quiescent and activated stellate cell groups







## Supplementary Table 9: Use of Fewer Genes in the Gene Signature Improves Outcomes Association in the Resected HCC and Hepatitis C Cirrhosis Cohorts

			# signature genes used					
			122	81	55	31	17	6
HCV Cirrhosis Cohort	Survival	NES	1.655	2.188	2.797	<u>3.048</u>	2.485	2.107
		FDR q-val	0.014	<u>&lt;0.001</u>	<u>&lt;0.001</u>	<u>&lt;0.001</u>	<u>&lt;0.001</u>	0.001
	Child-Pugh class progression	NES	1.704	2.475	2.953	<u>3.258</u>	2.873	2.001
		FDR q-val	0.014	<u>&lt;0.001</u>	<u>&lt;0.001</u>	<u>&lt;0.001</u>	<u>&lt;0.001</u>	<u>&lt;0.001</u>
	Decompensation	NES	1.705	2.369	2.669	<u>2.974</u>	2.708	1.929
		FDR q-val	0.009	<u>&lt;0.001</u>	<u>&lt;0.001</u>	<u>&lt;0.001</u>	<u>&lt;0.001</u>	0.003
	HCC incidence	NES	1.199	1.905	2.103	<u>2.882</u>	2.578	2.024
		FDR q-val	0.221	0.003	<u>&lt;0.001</u>	<u>&lt;0.001</u>	<u>&lt;0.001</u>	0.002
Resected HCC Cohort	Survival	NES	1.971	2.153	<u>2.47</u>	1.62	1.989	1.662
		FDR q-val	0.001	<u>&lt;0.001</u>	<u>&lt;0.001</u>	0.028	0.001	0.023
	HCC recurrence	NES	2.169	2.526	<u>2.739</u>	2.079	1.808	1.792
		FDR q-val	<u>&lt;0.001</u>	<u>&lt;0.001</u>	<u>&lt;0.001</u>	0.001	0.011	0.01

NES: normalized enrichment score (GSEA, poor versus good outcomes)

FDR q-val: false discovery rate q-value (GSEA, poor versus good outcomes)

### Association of reduced gene signatures with survival in prognostic validation set (n=82, log-rank)

	HR	95%CI	p-value
55-gene hepatic stellate signature: positive (vs. negative)	2.78	1.28 – 6.06	0.007
31-gene hepatic stellate signature: positive (vs. negative)	2.06	0.96 – 4.39	0.055